Multi-input stochastic prediction of insulin sensitivity for tight glycaemic control using insulin sensitivity and blood glucose data

Shaun Davidson[1], Chris Pretty[1], Vincent Uyttendaele[2], Jennifer Knopp[1], Thomas Desaive[2], J. Geoffrey Chase[1]

[1]Department of Mechanical Engineering, University of Canterbury, Christchurch, New Zealand

[2]GIGA-Cardiovascular Sciences, University of Liège, Liège, Belgium

Phone: +64 3-369 0065

Fax: +64 3-369 0065

E-mail: shaun.davidson@canterbury.ac.nz

Address: 20 Kirkwood Ave, Upper Riccarton, Christchurch 8041

Abstract

Background: Glycaemic control in the intensive care unit is dependent on effective prediction of patient insulin sensitivity (*SI*). The stochastic targeted (STAR) controller uses a 2D stochastic model for prediction, with current *SI* as an input and future *SI* as an output.

Methods: This paper develops an extension of the STAR 2D stochastic model into 3D by adding blood glucose (*G*) as an input. The performance of the 2D and 3D stochastic models is compared over a retrospective cohort of 65,269 data points across 1,525 patients.

Results: Under five-fold cross-validation, the 3D model was found to better match the expected potion of data points within, above and below various credible intervals, suggesting it provided a better representation of the underlying probability field. The 3D model was also found to provide an 18.1% narrower 90% credible interval on average, and a narrower 90% credible interval in 96.4% of cases, suggesting it provided more accurate predictions of future *SI*. Additionally, the 3D stochastic model was found to avoid the undesirable tendency of the 2D model to overestimate *SI* for patients with high *G*, and underestimate *SI* for patients with low *G*.

Conclusions: Overall, the 3D stochastic model is shown to provide clear potential benefits over the 2D model for minimal clinical cost or effort, though further exploration into whether these improvements in *SI* prediction translate into improved clinical outcomes is required.

Keywords

Glycaemic Control; Stochastic Model; Gaussian Kernel; Insulin Sensitivity; Stochastic Targeted

1 Introduction

Stress-induced hyperglycaemia is frequently observed in intensive care unit (ICU) patients [1-3]. Both this hyperglycaemia [1, 2, 4] and general glycaemic variability [5, 6] have been associated with increased morbidity and mortality. As such, effective glycaemic control that is capable of regulating blood glucose (*G*) levels to within a safe and relatively narrow band by modulating insulin and nutrition infusion rates may help alleviate these negative effects [7-12], improving patient outcomes and reducing cost of care.

However, patient insulin sensitivity (*SI*), which governs glycaemic response, undergoes complex and highly variable stress-induced changes [13, 14]. Thus, safe, effective control of patient *G* is reliant on effective and accurate prediction of future patient *SI* [15-17]. Failure to effectively account for variability in future *SI* can lead to an increased risk of hypoglycaemia, as in [18-25], where 17–29 % of patients had at least one blood glucose reading of less than 2.2 mmol/L. Hypoglycaemia is associated with increased mortality [26-28], and often occurs due to large, difficult to predict changes in patient *SI* over short periods [14, 29, 30]. These large changes in patient *SI* are particularly frequent in the first 48 hours post ICU admission [31], where there is a concordant, high rate of hypoglycaemic events and a strong association with mortality [27].

The stochastic targeted (STAR) protocol [15, 16] combines a clinically validated model of the insulin-glucose system [32, 33] with a stochastic model for predicting future variability in *SI* [34, 35]. STAR titrates patient and condition specific insulin and nutrition interventions, designed to minimise hyperglycaemia, while maintaining a low risk of a mild hypoglycaemia (no more than 5% risk of *G* less than 4.4 mmol/L [15-17]). Thus, STAR is designed to provide safe, effective, patient-specific glycaemic control, directly accounting for intra- and inter-

patient variability. Since 2011, STAR has been the standard of care in the ICUs of both the Christchurch Hospital, New Zealand, and Kálmán Pándy Hospital, Hungary, demonstrating safe and effective control to the targeted range in these two cohorts [36].

The current stochastic model prediction of future *SI* involves a 2D kernel combining current insulin sensitivity (SI_t) and future insulin sensitivity (SI_{t+1}) over a cohort of representative, retrospective data [34]. With the increased availability of retrospective information since STAR's inception in 2011, there has been interest in adding additional information and dimensions to this stochastic model. Previous studies have explored using further historic *SI* values (e.g. SI_{t-1}) or historic variability in *SI* [37, 38]. This study instead explores the effects of the addition of a different data type, current blood glucose (G_t), which is consistently measured in the ICU and is used in the ICING model as well as by the STAR protocol to titrate treatment, but not in the kernel-based model predictor. The overall aim of this paper is to develop a stochastic model to forward predict future SI_{t+1}, using current SI_t and current G_t as model inputs.

2 Materials and Methods

2.1 The ICING Model

STAR employs the clinically validated ICING metabolic model [35, 39-41] to account for dynamic *SI* as encountered in critical care. This model is defined:

$$\dot{G} = -p_G \cdot G - SI \cdot G \cdot \frac{Q}{1 + \alpha_G \cdot Q} + \frac{\min(d_2 \cdot P_2, P_{\max}) + EGP - CNS + PN(t)}{V_G} \quad (1)$$

$$\dot{I} = -\frac{n_L \cdot I}{1 + \alpha_I \cdot I} - n_K \cdot I - n_I \cdot (I - Q) + \frac{u_{ex}(t)}{V_I} + (1 - x_L)\frac{u_{en}(G)}{V_I} \quad (2)$$

$$\dot{Q} = n_I \cdot (I - Q) - n_C \cdot \frac{Q}{1 + \alpha_G \cdot Q} \quad (3)$$

$$\dot{P}_1 = -d_1 \cdot P_1 + P(t) \quad (4)$$

$$\dot{P}_2 = -\min(d_2 \cdot P_2, P_{\max}) + d_1 \cdot P_1 \quad (5)$$

where G (mmol/L) is the blood glucose concentration, I (mU/L) is the blood insulin concentration and Q (mU/L) is the interstitial insulin concentration. P_1 and P_2 represent the total stomach and gut glucose, respectively, and serve to model the rate at which enteral feed $P(t)$ (mmol/min) enters the blood. Clinically measured or recorded model inputs are enteral feed rate $P(t)$, parenteral feed rate $PN(t)$, blood glucose $G(t)$, and exogenous insulin infusion rate $u_{ex}(t)$. Pancreatic insulin secretion, $u_{en}(G)$, is modelled as a linear piecewise function of $G(t)$ [41]. The remaining parameters are fixed, population standard values, and are provided in Table 1. Integral based fitting [42, 43] is used to fit a single parameter: a constant *SI* over a period of 60 minutes. To fit this *SI*, a linear profile for $G(t)$ is assumed, as well as steady state initial conditions if no previous I or Q has been calculated for the patient (i.e. for the initial measurement for each patient). Further details on the ICING model can be found in [43, 44].

Table 1: Population standard parameter values for the ICING metabolic model

Parameter	*Value*	*Description*
p_G	0.006 min^{-1}	Endogenous glucose clearance
α_G	1/65 l/mU	Saturation of insulin-mediated glucose uptake
V_G	13.3 L	Glucose distribution volume
EGP	1.16 mmol/min	Endogenous glucose production (hepatic)
CNS	0.3 mmol/min	Glucose uptake by central nervous system
d_1	-ln(0.5)/20 min^{-1}	Gastric emptying of stomach to gut
d_2	-ln(0.5)/100 min^{-1}	Glucose absorption for gut to bloodstream
P_{max}	6.11 mmol/min	Maximum glucose absorption rate from gut
x_L	0.67	Fractional first pass hepatic insulin clearance from portal vein
n_L	0.1578 min^{-1}	General hepatic insulin clearance
α_I	1.7 x 10^{-3} L/mU	Saturation of hepatic insulin clearance
n_K	0.0542 min^{-1}	Kidney clearance of insulin
n_C	0.0075 min^{-1}	Cellular degradation of internalised insulin
n_I	0.0075 min^{-1}	Insulin diffusion between plasma and interstitium
k_1	14.9 mU·L/mmol/min	Insulin secretion model parameter
k_2	-49.9 mU/min	Insulin secretion model parameter
u_{min}	16.7 mU/min	Minimal insulin secretion
u_{max}	266.7 mU/min	Maximal insulin secretion
V_I	4.0 L	Insulin distribution volume

The *SI* parameter in the ICING model has been shown to broadly capture patient metabolic response to insulin and nutrition inputs, and therefore be a key factor in control efficacy [29]. Thus, *SI* provided by the ICING model is both physiologically relevant and important to the safety and efficacy of glycaemic control.

2.2 Clinical Data Set

2.2.1 Cohort

The data employed in this study is drawn from three cohorts also employed in [36], largely gathered using the STAR tablet interface:

1. Patients from the Christchurch Hospital ICU, Christchurch, New Zealand, treated using the STAR protocol between June 2011 and May 2015.

2. Patients from the Kálmán Pándy Hospital ICU, Gyula, Hungary, treated using the STAR protocol between December 2011 and May 2015.
3. Patients from the Christchurch Hospital ICU, Christchurch, NZ, treated using the Specialized Relative Insulin and Nutrition Tables (SPRINT) protocol [8, 10], the predecessor of STAR, between July 2005 and May 2007.

And one additional cohort:

4. Patients from the Christchurch Hospital ICU, Christchurch, New Zealand, treated using the STAR protocol and a modified, clinically driven nutrition protocol between June 2015 and November 2017.

Unlike in [36], where the goal was the assessment of the effectiveness of the STAR controller, and patients who were treated outside of guidelines or for a short period of time were excluded, the focus here is to capture glucose-insulin dynamics, which can be captured irrespective of compliance to the protocol. Overall, this data set encompassed 1,525 patients. For information on the demographics of cohorts 1 – 3, see [36].

2.2.2 Ethics, consent and permissions

Approval for a retrospective audit, analysis and publication of Christchurch patient data was provided by the Upper South Regional Ethics Committee, New Zealand. According to local ethical codes in Hungary, publication of anonymized, retrospective data, as here, is considered a clinical data audit.

2.2.3 Data Processing

The patient data set included enteral feed rate $P(t)$, parenteral feed rate $PN(t)$, exogenous insulin infusion rate $u_{ex}(t)$, and plasma glucose $G(t)$ measurements at intervals of 1 to 3 hours. $G(t)$

measurement was performed using an Arkray Super-Glucocard™ II glucometer (Arkray, Minnesota, USA) (2011–2012) or a Roche Accu-Chek Inform II (F. Hoffmann-La Roche Ltd, Basel, Switzerland) (2012–2017) and blood taken directly from an arterial line. At each glucose assay point, SI_t was derived by solving for a constant *SI* over the 60-minute period directly prior to the assay and SI_{t+1} by solving for a constant *SI* over the 60-minute period directly after the assay [42, 45]. As such, only patient data points with both a prior and future G_t assay available were usable in the model, giving a total of 65,269 usable patient data points.

As shown in Fig. 1, the overall distribution of the datasets SI_t, G_t and SI_{t+1} all appear to be log-normal, and *SI* and *G*, for physiological and physical reasons, should not be negative, corresponding to the logarithmic domain. As several kernel rules employed in this paper (e.g. Silverman's Rule [46]) are defined under assumptions of normally distributed data, all three variables were log transformed prior to generation of the 3D kernel field. This transformation shifts the log-normal data into an apparently normal distribution, as shown in Fig. 1, though the data set is too large for normality tests to be performed effectively [47].

As demonstrated by the effectiveness of the current 2D stochastic model [34], there are strong correlations between the data even prior to this transformation. Stochastic prediction of future *SI* using current *SI* is supported by a strong autocorrelation, with values of $R = 0.87$, 0.85, and 0.80 at lags of $t = 1$, 2, and 3 hours respectively, where 3 hours is the maximum prediction interval used in the STAR protocol [15, 16]. These values were calculated by using non-transformed *SI*, as in the 2D stochastic model, and normalising for mean and standard deviation.

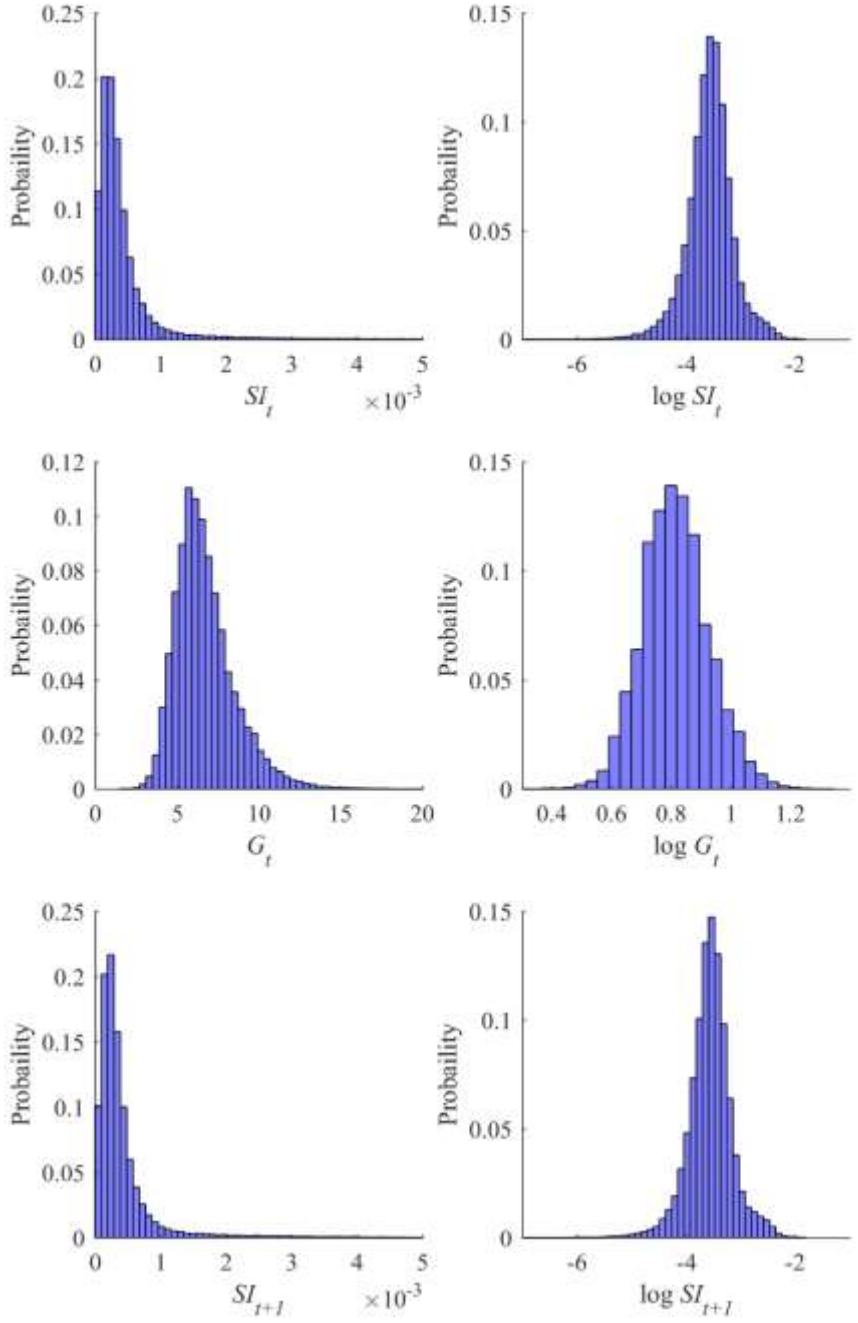

Fig. 1: Normal and log transformed data distributions for SI_t, G_t and SI_{t+1}, 65,269 data points.

2.3 3D Stochastic Model

This model employs a multivariate Gaussian normal kernel, which will behave in a log-normal fashion due to the logarithmic transformations on the input data. For *n* variables, the Gaussian normal kernel can be generally defined:

$$f(\mathbf{y}) = \frac{\exp\left(-\frac{1}{2}(\mathbf{y}-\mathbf{x})^{\mathrm{T}}\mathbf{C}^{-1}(\mathbf{y}-\mathbf{x})\right)}{\sqrt{(2\pi)^k|\mathbf{C}|}} \tag{6}$$

Where **C** is the *n* x *n* covariance matrix, **y** is a query point in *n*-dimensional space, and **x** is the centre point of the kernel.

2.3.1 Ortho-Normalisation

Ortho-normalisation involves transformation of the data so that it is centred at 0, has a standard deviation and variance of 1 in every direction, and no correlation exists between the transformed variables. These properties have a number of positive effects on kernel calculation, as will be shown.

Define an array **X** of *k* kernel data points across these *n* variables:

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & & \vdots \\ x_{k,n} & \cdots & x_{k,n} \end{bmatrix} \tag{7}$$

Perform the Cholesky decomposition on **C**, the *n* x *n* covariance matrix between the variables in **X**, and use the resulting matrix **R** to define the *n* x *n* transformation matrix **A**:

$$\mathbf{C} = \mathbf{R} \times \mathbf{R}^{\mathrm{T}} \tag{8a}$$

$$\mathbf{A} = (\mathbf{R}^{\mathrm{T}})^{-1} \tag{8b}$$

Adjust each of the *n* variables in **X** to have a mean of 0:

$$\widehat{\mathbf{X}} = \mathbf{X} - \mathbf{e}_k \times \bar{\mathbf{x}} \tag{9}$$

Where $\mathbf{e}_k$ is a k x 1 vector of ones and $\bar{\mathbf{x}}$ is a 1 x n vector containing the mean value of each of the n variables. The zero-meaned variables can now be transformed into ortho-normalised space:

$$\widetilde{\mathbf{X}} = \mathbf{A} \times \widehat{\mathbf{X}} \tag{10}$$

Given the aforementioned properties of ortho-normalised space, the ortho-normalised covariance matrix $\tilde{\mathbf{C}}$ is the identity matrix. As such, the iso-probability shells of the Gaussian normal kernels will take the form of n-dimensional hyperspheres in this ortho-normalised space, while behaving as non-axis aligned n-dimensional hyperellipsoids in solution space. An example of this in 2D is shown in Fig. 2.

In ortho-normalised space, the kernel equation (Eq. 6) can be significantly simplified to:

$$f(\mathbf{y}) = \frac{\exp\left(-\frac{1}{2}\left(\frac{\sum_{i=1}^{n}(\tilde{y}_i - \tilde{x}_i)^2}{\tilde{\sigma}^2}\right)\right)}{(2\pi)^{\frac{n}{2}}\tilde{\sigma}^n} \tag{11}$$

Where:

$$\tilde{\mathbf{y}} = \mathbf{A} \times \ (\mathbf{y} - \bar{\mathbf{x}}) \tag{12a}$$

$$\tilde{\mathbf{x}} = \mathbf{A} \times \ (\mathbf{x} - \bar{\mathbf{x}}) \tag{12b}$$

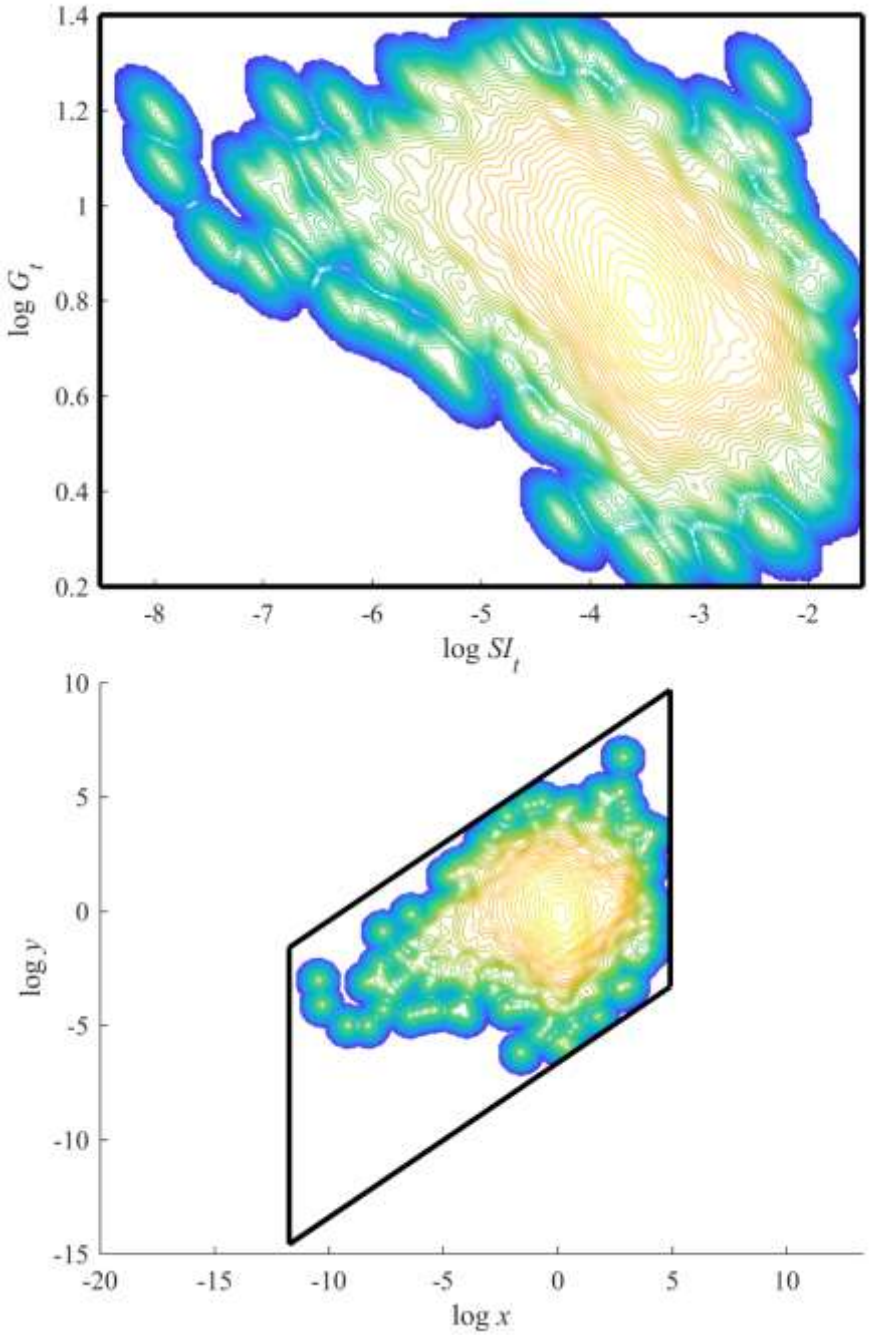

Fig. 2: The rectangular grid in solution space (top) becomes a parallelogram in ortho-normalised space (bottom), causing the circular kernels in ortho-normalised space (bottom) to appear as non-axis aligned ellipses in solution space (top). Note the contours are log-density contours, selected to highlight the boundaries of individual kernels, and that the outlying kernels would not appear in a normal density plot.

2.3.2 Kernel Specific Variance, $\tilde{\sigma}$

For each ortho-normalised kernel data point, $\tilde{\mathbf{x}}_j$, in the matrix $\tilde{\mathbf{X}}$, there is an associated kernel variance, $\tilde{\sigma}_j$. Note that, as the iso-probability shells of the ortho-normalised kernels are spherical, this $\tilde{\sigma}_j$ applies in all directions for a given kernel data point. Combining Silverman's rule of thumb with a relative local density factor as in [34] yields Eq. 13 for $\tilde{\sigma}_j$ in the 3D case (full derivation in Appendix A):

$$\tilde{\sigma}_j = 0.969 \times \left(\frac{m_j {R_Z}^3 k^{\frac{3}{7}}}{Z} \right)^{-\frac{1}{7}} \tag{13}$$

for a given ortho-normalised kernel data point $\tilde{\mathbf{x}}_j$. Here m_j is the number of points in $\tilde{\mathbf{X}}$ within a radius $k^{-\frac{1}{7}}$ from the kernel data point $\tilde{\mathbf{x}}_j$ and R_Z is the radius from the ortho-normalised centre point of 0 that contains $Z \cdot k$ kernel data points for some $Z \leq 1$. Z allows for the exclusion of outlying data points that can dramatically alter the global data density calculation.

2.3.3 Sampling Probability Density Functions for Patient Data Points

At a query point, $\mathbf{y}$, the probability density, $f(\mathbf{y})$, is the sum of all k kernel data point probability densities, divided by global sum of these probability densities, k. In 3D:

$$f(\mathbf{y}) = \frac{1}{k} \sum_{j=1}^{k} \left\{ \frac{\exp\left(-\frac{1}{2} \left(\frac{\sum_{i=1}^{3} \left(\tilde{y}_i - \tilde{x}_{j,i} \right)^2}{{\tilde{\sigma}_j}^2} \right) \right)}{(2\pi)^{\frac{3}{2}} {\tilde{\sigma}_j}^3} \right\} \tag{14}$$

where:

$$\tilde{\mathbf{y}} = \mathbf{A} \times \left(\mathbf{y} - \bar{\mathbf{x}} \right) \tag{15a}$$

$$\tilde{\mathbf{x}}_j = \mathbf{A} \times \left(\mathbf{x}_j - \bar{\mathbf{x}}\right) \tag{15b}$$

Here, there are 2 pieces of input patient data, $y_1 = SI_t$ and $y_2 = G_t$, and 1 output for which a probability density function (PDF) is sought, $y_3 = SI_{t+1}$. Determining the conditional 1D PDF, f_1, for a given patient data point $\mathbf{y}_{1:2}$ involves sampling across the range of potential values for y_3, typically from $SI_{t+1,\min} = 10^{-8.5}$ L/mU/min to $SI_{t+1,\max} = 10^{-1.5}$ L/mU/min, for the given $\mathbf{y}_{1:2}$ and normalising:

$$f_1(\mathbf{Y}) = \frac{f(\mathbf{Y})}{\int_{y_3=SI_{t+1,\min}}^{SI_{t+1,\max}} f(\mathbf{Y})\,\mathrm{d}y_3} \tag{16}$$

where $f(\mathbf{Y})$ refers specifically to Eq. 14, and $\mathbf{Y}$ is the line consisting of points $\mathbf{y}$ across the full range of possible y_3 values at $\mathbf{y}_{1:2}$:

$$SI_{t+1,\min} \leq y_3 \leq SI_{t+1,\max} \tag{17}$$

The first 2 transformed components of the query point $\tilde{\mathbf{y}}_{1:2}$ and kernel centre point $\tilde{\mathbf{x}}_{j,1:2}$ are independent of changes in y_3 due to the transformation matrix $\mathbf{A}$ being upper triangular. This means Eq. 14 can be re-written:

$$f(\mathbf{y}) = \frac{1}{k}\sum_{j=1}^{k}\left\{\frac{\exp\left(-\frac{1}{2}\left(\frac{\sum_{i=1}^{2}\left(\tilde{y}_i - \tilde{x}_{j,i}\right)^2}{\tilde{\sigma}_j{}^2}\right)\right)}{(2\pi)^{\frac{3}{2}}\tilde{\sigma}_j{}^3} \times \exp\left(-\frac{1}{2}\left(\frac{\left(\tilde{y}_3 - \tilde{x}_{j,3}\right)^2}{\tilde{\sigma}_j{}^2}\right)\right)\right\} \tag{18}$$

$$SI_{t+1,\min} \leq y_3 \leq SI_{t+1,\max} \tag{19}$$

This modification is an application of conditional probability rules, which can be used here as the transformed variables in $\tilde{\mathbf{y}}$ and $\tilde{\mathbf{x}}$ are independent and uncorrelated (unlike their non-transformed counterparts, another advantage of ortho-normalisation). The first term in Eq. 19 remains constant across the entire sample line $\mathbf{Y}$ as it is only dependent on $\tilde{y}_1$ and $\tilde{y}_2$, and thus

only needs to be calculated once per kernel centre point $\tilde{\mathbf{x}}_j$. The second term, which includes $\tilde{y}_3$, is the only part that varies, for a given kernel centre point, along the line **Y**.

Further computational efficiency can be gained when calculating $f_1(\mathbf{Y})$ by restricting calculation of constituent kernel probability density functions to cases where, for a given sample point, **y** on the line **Y**:

$$\left\|\tilde{\mathbf{y}} - \tilde{\mathbf{x}}_j\right\|_2 < 5\tilde{\sigma}_j \tag{20}$$

Given the spherical shape of kernel iso-probability shells in 3D ortho-normalised space ensures $\tilde{\sigma}_j$ is identical in all directions, this restriction can easily be set geometrically, without requiring the calculation of a norm at every point **y**, as explained in Appendix B. The selected radius can be varied based on accuracy and computational load requirements. A radius of $5\tilde{\sigma}_j$ contains 99.998% of the probability density function in 3D, compared to 99.887% at $4\tilde{\sigma}_j$ and 97.071% at $3\tilde{\sigma}_j$. 1D PDFs were found using pre-calculated values for the transformation matrix **A** and $\tilde{\sigma}_j$ values, neither of which change throughout sample space, and at a resolution of 300 points along the PDF. Under these conditions, the alteration in Eq. 20 decreased computational time from 28.38 s per PDF, using conditional probabilities as in Eq. 18, to 0.24s, using both Eqs. 18 and 20, on average, a factor of over 118 times.

STAR currently uses a 400 point table of SI_t values with pre-calculated 5% and 95% SI_{t+1} credible interval bounds, and interpolates to find the credible interval bounds for a given patient SI_t. However, this approach is considerably less desirable for a 3D kernel field with a 2D input. First, for a 2D input there is a need to store a 400 by 400 table of 160,000 pre-calculated 5% and 95% bounds to achieve the same resolution as the 1D table. Second, interpolation error increases in higher dimensions due to interactions between errors in individual input variables. Third, the contours of the probability density field can be extremely steep, especially with log-

normal kernels, making interpolation generally less desirable. The improvement in computational efficiency provided by Eq. 20 potentially allows the 'live' calculation of credible interval boundaries, the equivalent of an infinite grid resolution, addressing the various concerns listed above. In addition, this approach provides the entire PDF rather than solely the 5% and 95% bounds. Finally, 'live' calculation of credible interval boundaries raises the possibility of real-time updating of the stochastic model using weighted kernels for the current patient's previous data points, increasing model personalisation.

2.4 2D Stochastic Model

The existing stochastic model employs a 2D multivariate Gaussian normal kernel with SI_t used as an input for prediction of SI_{t+1}. The primary differences between this model and the 3D model, other than the shift between 3D and 2D, are:

- The data are not log transformed. Thus, the kernels have a normal, rather than log-normal, form. As this difference allows kernel mass to exist below $SI = 0$, any kernel mass below 0 is removed, and each kernel then normalised to restore that kernel's mass to 1.0.
- Covariance terms between SI_t and SI_{t+1} are not considered, thus the major and minor axes of the kernel iso-density contours are aligned with the axes SI_t and SI_{t+1}. Accordingly, the data does not undergo an ortho-normalisation, and for each kernel σ is calculated separately for SI_t and SI_{t+1}.

The kernel equation for a sample point $\mathbf{y}$ in the 2D case defined:

$$f(\mathbf{y}) = \frac{1}{k}\sum_{j=1}^{k}\left\{\frac{\exp\left(-\frac{1}{2}\sum_{i=1}^{2}\left(\frac{\left(y_i - x_{j,i}\right)^2}{{\sigma_{j,i}}^2}\right)\right)}{2\pi\sqrt{\prod_{i=1}^{2}\sigma_{j,i}}}\right\} \tag{21}$$

where kernel and axis specific variance is calculated via:

$$\sigma_{j,i} = \left(\frac{m_j {R_Z}^2 k^{\frac{1}{3}}}{Z}\right)^{-\frac{1}{6}} \times D_i \tag{22}$$

Here, D_i is the global standard deviation for the i^{th} variable. Full details on this stochastic model can be found in [34, 35].

2.5 Validation and Assessment

Several analyses were performed to validate the model was functioning correctly and evaluate the resultant performance changes from the existing 2D model. Note both models were built across the entire, current cohort of patient data for equivalence.

2.5.1 Five-Fold Cross-Validation

To mitigate issues with developing and then validating the data over the same data set, across which, by definition, each credible interval would contain the correct amount of data, five-fold cross-validation was used [48]. Cross-validation involved breaking the data into 5 approximately equally sized sets, and testing each data set on a model built from the other 4 data sets. This process ensures that the model is built and then assessed on separate datasets each time.

Slight variation in data set sizes arose due to the decision to treat each patient's set of data points as a discrete block when dividing the data, such that every patient data point was tested

on a model developed without any data from that same patient. This analysis allows evaluation of the accuracy of the credible intervals generated by the model, and thus whether an appropriate kernel width, $\tilde{\sigma}$, has been selected. Kernel width was modified by altering the term Z, changing the normalisation component of the relative data density term in Eq. 13. The final values selected were $Z = 0.97$ for the 2D model and $Z = 0.999$ for the 3D model, based on cross-validation results with the goal of providing as accurate a 90% credible interval as possible.

2.5.2 Comparison of Credible interval Bounds

This analysis involves a comparison of the credible interval boundaries and widths across all patient data, using cross validation to avoid building and evaluating a model on the same data set. It intends to provide a general indication and summary statistics of the confidence with which estimation of SI_{t+1} can be achieved, using SI_t and G_t. Particular focus is given to the 90% credible interval boundary, as the 5^{th} and 95^{th} percentiles are used by STAR to titrate treatment [15, 16]. This analysis was accompanied by a visual representation of the credible interval boundary surfaces between the 3D and 2D models, built using the entire data set, allowing visual comparison of trends in these boundary surfaces across the two models.

2.5.3 Distribution of Data Points Outside Bounds

An analysis of the distribution of data points that sat above or below the 90% credible interval, the interval boundary used clinically by STAR was carried out as this cohort of patients are those for whom the model had over- or under- estimated *SI*. The aim is to examine the clinical risk of model over- or under- estimation in these patients.

3 Results

3.1 Five-Fold Cross-Validation

The percentage of SI_{t+1} data points within, above, and below forward credible interval bounds for the 2D and 3D stochastic models using the five sets of cross-validated data are shown in Table 2. The percentage of SI_{t+1} data points within, above, and below bounds for the 3D model correspond more closely to the expected percentages for all intervals assessed, except for the percentage of data points within the 90% credible interval (90.28% for the 3D kernel and 90.24% for the 2D kernel). Across the bounds tested, the 3D model provides percentages of data points within, above and below bounds that are 2.6 times closer to the expected percentages, on average, than the 2D model. Additionally, at the 90% and 95% credible interval for the 2D model there is a notably greater portion of points below than above the credible interval bounds, while the 3D model maintains reasonably even portions of both. Further, the 2D model fails to provide 95% of points within the 95% credible interval bounds, while the 3D model continues to show a good match to the data. Thus, the 3D model is more reflective of the underlying data than the 2D model, and may be better suited to predict SI_{t+1}.

Table 2: Cross-Validation Results, 2D and 3D stochastic models. Values reflect the percentage of SI_{t+1} points that fell below, within, or above credible interval bounds.

Credible Interval	2D Stochastic			3D Stochastic		
	Below	Within	Above	Below	Within	Above
95%	3.77%	94.06%	2.16%	2.50%	95.06%	2.43%
90%	5.45%	90.24%	4.29%	4.92%	90.28%	4.80%
75%	11.12%	77.96%	10.91%	11.93%	76.28%	11.78%
50%	22.53%	54.60%	22.86%	23.95%	52.02%	24.02%

3.2 Credible Interval Bounds

Credible interval summary statistics for SI_{t+1} using the 2D and 3D models are presented in Table 3. These statistics were generated by calculating credible intervals for each patient data point using the previously defined cross-validation model that excluded that patient, and then taking the average across all patient data points. From Table 3, the 3D 90% credible interval is 18.12% narrower on average, with a 6.37% lower 95% boundary and a 14.38% higher 5% boundary. Across the full set of data points, the 3D model provides a narrower 90% credible interval in 96.15% of cases, and is narrower by 20.95% on average in these cases. The 3D model provides a wider credible interval in 3.85% of cases, where it is wider by an average amount of 21.21%. Where the 3D model is wider, the mean G_t is 6.02 mmol/L, less than the overall average of 6.74 mmol/L, and the mean SI_t is 8.75 x 10^{-4} L/mU/min, over double the overall average of 4.24 x 10^{-4} L/mU/min. Thus, the 3D model is more conservative where risk is higher, and narrower where stability in *SI* allows improved prediction.

Table 3: Credible interval Summary Statistics. These figures are averages of credible intervals generated using a five-fold cross-validation model for each patient data point.

Model	2D Stochastic	3D Stochastic
Expected Value	4.24 x 10^{-4} L/mU/min	4.24 x 10^{-4} L/mU/min
5% Boundary	2.26 x 10^{-4} L/mU/min	2.59 x 10^{-4} L/mU/min
95% Boundary	6.25 x 10^{-4} L/mU/min	5.85 x 10^{-4} L/mU/min
90% Interval Width	3.99 x 10^{-4} L/mU/min	3.27 x 10^{-4} L/mU/min

Figures 3 and 4 show the 5% and 95% boundary surfaces for the 2D and 3D stochastic models. Overlaid upon these figures is the density of clinical data, providing emphasis on which regions of the surface are most likely to be encountered clinically.

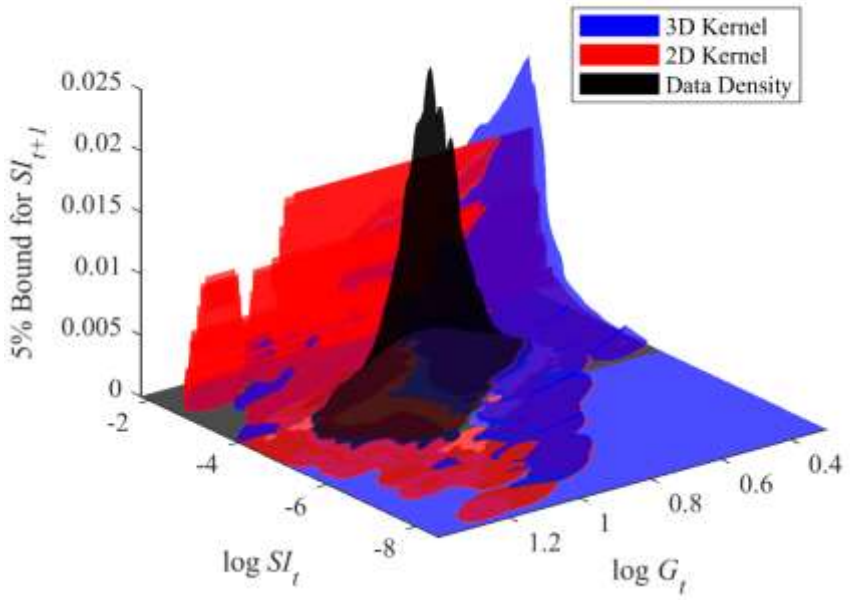

Fig. 3: The 5% bound for SI_{t+1} (higher is generally better) for the 2D and 3D stochastic models, with data density overlaid

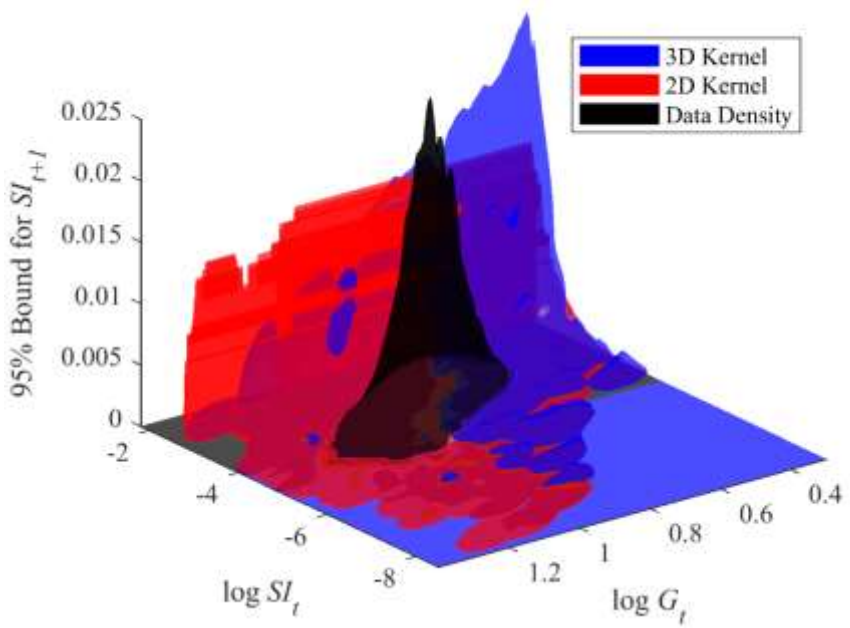

Fig. 4: The 95% bound for SI_{t+1} (lower is generally better) for the 2D and 3D stochastic models, with data density overlaid

Of note in Figures 3 and 4 is the tendency for the 3D kernel surface to be higher for lower values of G_t, and lower for high values of G_t, showing a clear dependence missed by the 2D model. Figure 5 shows the 90% interval width for the 2D and 3D stochastic models. Here it is clear that the 3D model produces a narrower credible interval in the majority of cases, especially at common clinical values.

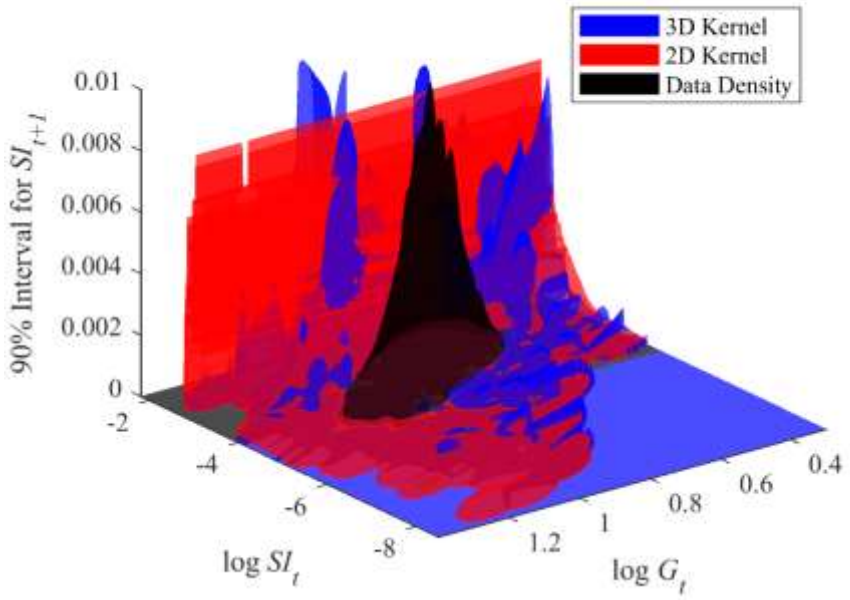

Fig. 5: The 90% credible interval width for SI_{t+1} (narrower is better) for the 2D and 3D stochastic models, with data density overlaid

3.3 Data Points Outside Bounds

Fig 6. shows histograms of the distribution of G_t, SI_t and SI_{t+1} data point sets within, above, and below the 90% cross-validated credible interval for the 2D and 3D models. Of interest is the bimodal distribution in the 'below bounds' data set from the 2D model for SI_t and SI_{t+1}. Also of note is the tendency of the 2D model to overestimate *SI* of patients with high G_t, and underestimate *SI* of patients with low G_t.

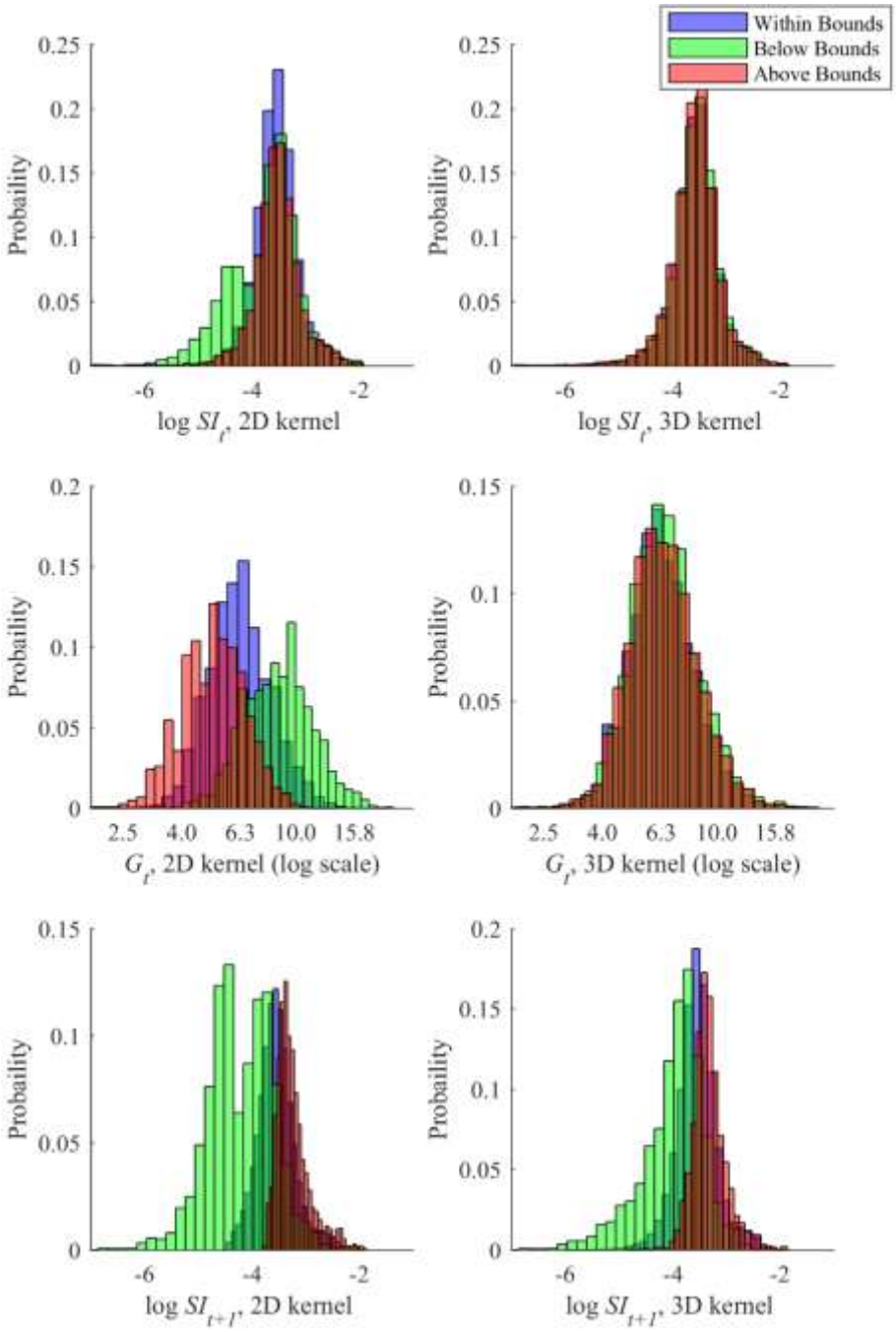

Fig. 6: Histograms showing the data distribution of points within, above and below the 90% credible interval bounds for the 2D and 3D stochastic models.

4 Discussion

The cross-validation results in Table 2 show that the 3D stochastic model provides a better representation of the underlying data distribution than the 2D stochastic model, resulting in improved prediction of SI_{t+1}. On average, the 3D model provides percentages of predicted SI_{t+1} points within, above and below credible interval bounds that are 2.6 times closer to their expected percentages than the 2D model. Further, the difference between the expected and observed percentage of points within bounds varies less for the 3D model (0.06% to 2.02%) than for the 2D model (-0.94% to 4.60%), suggested the log transformation of the data and addition of G_t as an axis have resulted in a more representative kernel field. Finally, the 3D model maintains relatively even portions of points above and below the credible interval bounds, while for the 90% and 95% credible interval, the 2D model provides notably more points below than above the credible interval bounds. This result is likely due to the removal of kernel mass where $SI < 0$ for the 2D model, which shifts the centres of these kernels upwards and results in systematic overestimation of low values of SI_{t+1}.

These results are consistent across the repeated developing of the model on 80% of the data and evaluating it on the remaining, separate 20% of the data during cross-validation. This behaviour suggests that model performance is not a function of data selection, and thus the model and kernel choices are valid for the population. Further, the underlying data was gathered across multiple glycaemic control protocols, and multiple ICUs in different countries with different standards of clinical practice. Thus, the underlying data set is heterogenous, and the model and results would be expected to generalise well.

The results in Table 3 show that the 3D stochastic model is able to provide more accurate predictions of SI_{t+1} than the 2D stochastic model, enabling more targeted insulin-nutrition

treatments. The 3D model provides an 18.12% narrower 90% credible interval than the 2D model, on average, and a narrower credible interval in 96.15% of cases. This suggests the 2D model can be unnecessarily over-conservative in insulin dosing. As this analysis uses the cross validated data, with 90.24% of points within the credible interval for the 2D model and 90.28% of points within the credible interval for the 3D model, this improvement in performance is unlikely to be the result of overfitting. While this result is not entirely unsurprising, given the 3D model incorporates additional information unavailable to the 2D model, these results support that this additional information has been well selected and correctly implemented.

During the 3.85% of the time the credible interval is wider for the 3D model than the 2D model, G_t is lower than average (6.02 mmol/L vs 6.74 mmol/L) and SI_t is over double the average (8.75×10^{-4} L/mU/min vs 4.24×10^{-4} L/mU/min). This result suggests the 3D model is more conservative than the existing model for patients with very high SI_t and lower than average G_t, desirable to avoid hypoglycaemia. As such, use of more conservative credible intervals for this small subset of patients may well be beneficial and provide improved safety.

The credible interval surfaces in Figs. 3 – 5 suggest a general upward trend at lower levels of G_t for the prediction of SI_{t+1} using the 3D model, a trend which does not, by definition, exist for the 2D model. It is possible that this trend is representative of underestimation of SI_{t+1} leading to mild over-administration of insulin and a reduction in G_t. Thus, the 3D model serves to correct for over- or under- estimation of SI_{t+1} based on patient response.

The histograms in Fig. 6 provide a closer look at the patients that fall outside the credible interval boundaries for each model, compared to those that lie within the credible interval boundaries. The histograms of patient SI_t show that there is a bimodal peak to the left of overall mean SI_t in the ‘below bounds’ group for the 2D stochastic model, which does not exist in the

3D stochastic model. This bimodal peak is likely due to the aforementioned removal of kernel mass where $SI < 0$ in the 2D model, leading to systematic overestimation of low values of SI_{t+1}, a problem that is avoided by the log transformation of data in the 3D model, which in turn improves model shape and accuracy.

The histograms in Fig. 6 of patient G_t show that the 2D model, which is blind to G_t, tends to overestimate SI_{t+1} for patients with high G_t, increasing the risk of hyperglycaemia by anticipating these patients will be more responsive to insulin than they will actually be. Similarly, the 2D model tends to underestimate SI_{t+1} for patients with low G_t, increasing the risk of hypoglycaemia by anticipating these patients will be less responsive to insulin than they will actually be. The 3D model, which incorporates G_t, entirely removes these tendencies, which has the potential to have significant clinical benefits, as avoiding these hyper- and hypo-glycaemic cases is one of the overall goals of glycaemic control, where STAR already has excellent safety results [36]. Finally, the histograms of patient SI_{t+1} show that patients with higher than average SI_{t+1} are more likely to be underestimated, and patients with lower than average SI_{t+1} are more likely to be overestimated, which is to be expected. The bimodal peak for the 2D model present in SI_t is similarly present here.

While the data set employed is heterogenous, spanning multiple ICUs and glycaemic control protocols across a number of years, it still represents a small portion of the full range of patient populations and clinical standards of care. As such, validation of the model across additional patient data sets drawn from different ICUs with different protocols would add further robustness to the validation of the 3D model. Another important limitation of this study is the lack of data on actual effects on clinical hyper- and hypo- glycaemic incidents from using the 3D stochastic model for prediction, the ultimate end goal of any modifications to the STAR controller. Either a retrospective patient model of the full kernel model-controller-patient

interaction should be assessed, or experimental/clinical work done to assess whether the benefits provided here translate into improved clinical performance. Virtual trials [39] would be a first major step in showing improved utility, and this will constitute future work exploring if the narrowing of *SI* prediction ranges translates to appreciatively tighter and safer glycaemic control.

Overall, between the log transformation of the data and the addition of G_t as a data axis, the 3D model appears to better represent the probability field, provide more accurate predictions of SI_{t+1} and remove the causes for some undesirable behaviour in the 2D stochastic model. It is worth noting the 3D model does not require any data not currently gathered and used in the STAR protocol, and that, with the computational optimisations mentioned, the increase in computational load or runtime associated with adding this additional data axis is not significant on a clinical time scale. Thus, the improvements listed above can be provided for negligible computational or clinical cost and effort. While this study employs SPRINT and STAR data, the result should generalise to any glycaemic control protocol that relies on insulin sensitivity prediction.

5 Conclusions

A 3D stochastic model was developed for the prediction of future patient *SI* in the ICU. This model was compared to an existing, 2D stochastic model employed by the STAR controller over a retrospective cohort of 65,269 data points across 1,525 patients. The 3D stochastic model was shown to provide a better representation of the overall probability field, with a notably better conformation to expected percentage of data points within, above and below various credible interval widths while undergoing five-fold cross validation. The 3D stochastic model also provided more accurate predictions of *SI*, with an 18.12% narrower credible interval width on average, and a narrower credible interval width in 96.35% of cases. Finally, it was shown that the 2D stochastic model has a tendency to overestimate *SI* for patients with high blood glucose, and underestimate *SI* for patients with low blood glucose, increasing the risk of hyper- and hypo- glycaemia, respectively. In contrast, the 3D stochastic model was shown to completely avoid this tendency. As such, there is a clear case for further evaluation as to whether these improvements in predictive performance translate into an improvement in clinical hypo- and hyper- glycaemia rates when coupled with the STAR controller in an ICU.

Acknowledgements

The authors would like to acknowledge the contributions of Andrew Davidson to the development of the code, particularly with regards to optimisation, parallel processing, and data visualisation.

The authors would like to acknowledge funding from EU FP7, RSNZ IRSES mobility grants, Tertiary Education Commission of NZ (TEC) MedTech CoRE, NZ National Science Challenge 7 Science for Technological Innovation (SfTI) via Callaghan Innovation and MBIE, and the HRC of NZ.

References

1. Capes SE, Hunt D, Malmberg K, Gerstein HC: **Stress hyperglycaemia and increased risk of death after myocardial infarction in patients with and without diabetes: a systematic overview.** *The Lancet* 2000, **355:**773-778.
2. Finney SJ, Zekveld C, Elia A, Evans TW: **Glucose control and mortality in critically ill patients.** *JAMA* 2003, **290:**2041-2047.
3. McCowen KC, Malhotra A, Bistrian BR: **Stress-induced hyperglycemia.** *Crit Care Clin* 2001, **17:**107-124.
4. Krinsley JS: **Association between hyperglycemia and increased hospital mortality in a heterogeneous population of critically ill patients.** In *Mayo Clin Proc*. Elsevier; 2003: 1471-1478.
5. Egi M, Bellomo R, Stachowski E, French CJ, Hart G: **Variability of blood glucose concentration and short-term mortality in critically ill patients.** *The Journal of the American Society of Anesthesiologists* 2006, **105:**244-252.
6. Krinsley JS: **Glycemic variability: a strong independent predictor of mortality in critically ill patients.** *Crit Care Med* 2008, **36:**3008-3013.
7. Van Den Berghe G, Wouters P, Weekers F, Verwaest C, Bruyninckx F, Schetz M, Vlasselaers D, Ferdinande P, Lauwers P, Bouillon R: **Intensive insulin therapy in critically ill patients.** *N Engl J Med* 2001, **345:**1359-1367.
8. Chase JG, Shaw G, Le Compte A, Lonergan T, Willacy M, Wong X-W, Lin J, Lotz T, Lee D, Hann C: **Implementation and evaluation of the SPRINT protocol for tight glycaemic control in critically ill patients: a clinical practice change.** *Critical Care* 2008, **12:**R49.
9. Krinsley JS: **Effect of an intensive glucose management protocol on the mortality of critically ill adult patients.** In *Mayo Clin Proc*. Elsevier; 2004: 992-1000.
10. Chase JG, Pretty CG, Pfeifer L, Shaw GM, Preiser J-C, Le Compte AJ, Lin J, Hewett D, Moorhead KT, Desaive T: **Organ failure and tight glycemic control in the SPRINT study.** *Critical care* 2010, **14:**R154.
11. Krinsley JS, Jones RL: **Cost analysis of intensive glycemic control in critically ill adult patients.** *Chest* 2006, **129:**644-650.
12. Van den Berghe G, Wouters PJ, Kesteloot K, Hilleman DE: **Analysis of healthcare resource utilization with intensive insulin therapy in critically ill patients.** *Crit Care Med* 2006, **34:**612-616.
13. Dickson JL, Gunn CA, Chase JG: **Humans are horribly variable.** *Int J Clin Med Imaging* 2014, **1:**1-1000142.
14. Pretty C, Le Compte A, Chase JG, Shaw G, Preiser J-C, Penning S, Desaive T: **Variability of insulin sensitivity during the first 4 days of critical illness.** *Critical Care* 2012, **16:**P167.
15. Evans A, Le Compte A, Tan C-S, Ward L, Steel J, Pretty CG, Penning S, Suhaimi F, Shaw GM, Desaive T: **Stochastic targeted (STAR) glycemic control: design, safety, and performance.** *J Diabetes Sci Technol* 2012, **6:**102-115.
16. Evans A, Shaw GM, Le Compte A, Tan C-S, Ward L, Steel J, Pretty CG, Pfeifer L, Penning S, Suhaimi F: **Pilot proof of concept clinical trials of Stochastic Targeted (STAR) glycemic control.** *Annals of intensive care* 2011, **1:**38.

17. Fisk LM, Le Compte AJ, Shaw GM, Penning S, Desaive T, Chase JG: **STAR development and protocol comparison.** *IEEE Trans Biomed Eng* 2012, **59:**3357-3364.
18. Van den Berghe G, Wilmer A, Hermans G, Meersseman W, Wouters PJ, Milants I, Van Wijngaerden E, Bobbaers H, Bouillon R: **Intensive insulin therapy in the medical ICU.** *N Engl J Med* 2006, **354:**449-461.
19. Finfer S, Delaney A: **Tight glycemic control in critically ill adults.** *JAMA* 2008, **300:**963-965.
20. Brunkhorst FM, Engel C, Bloos F, Meier-Hellmann A, Ragaller M, Weiler N, Moerer O, Gruendling M, Oppert M, Grond S: **Intensive insulin therapy and pentastarch resuscitation in severe sepsis.** *N Engl J Med* 2008, **358:**125-139.
21. Griesdale DE, de Souza RJ, van Dam RM, Heyland DK, Cook DJ, Malhotra A, Dhaliwal R, Henderson WR, Chittock DR, Finfer S: **Intensive insulin therapy and mortality among critically ill patients: a meta-analysis including NICE-SUGAR study data.** *Can Med Assoc J* 2009, **180:**821-827.
22. Treggiari MM, Karir V, Yanez ND, Weiss NS, Daniel S, Deem SA: **Intensive insulin therapy and mortality in critically ill patients.** *Critical Care* 2008, **12:**R29.
23. Kalfon P, Giraudeau B, Ichai C, Guerrini A, Brechot N, Cinotti R, Dequin P-F, Riu-Poulenc B, Montravers P, Annane D: **Tight computerized versus conventional glucose control in the ICU: a randomized controlled trial.** *Intensive Care Med* 2014, **40:**171-181.
24. Hamimy W, Khedr H, Rushdi T, Zaghloul A, Hosni M, Aal AA: **Application of conventional blood glucose control strategy in surgical ICU in developing countries: Is it beneficial?** *Egyptian Journal of Anaesthesia* 2016, **32:**123-129.
25. Investigators N-SS: **Intensive versus conventional glucose control in critically ill patients.** *N Engl J Med* 2009, **360:**1283-1297.
26. Egi M, Bellomo R, Stachowski E, French CJ, Hart GK, Taori G, Hegarty C, Bailey M: **Hypoglycemia and outcome in critically ill patients.** In *Mayo Clin Proc*. Elsevier; 2010: 217-224.
27. Bagshaw SM, Bellomo R, Jacka MJ, Egi M, Hart GK, George C: **The impact of early hypoglycemia and blood glucose variability on outcome in critical illness.** *Critical care* 2009, **13:**R91.
28. Investigators N-SS: **Hypoglycemia and risk of death in critically ill patients.** *N Engl J Med* 2012, **367:**1108-1118.
29. Chase JG, Le Compte AJ, Suhaimi F, Shaw GM, Lynn A, Lin J, Pretty CG, Razak N, Parente JD, Hann CE: **Tight glycemic control in critical care–the leading role of insulin sensitivity and patient variability: a review and model-based analysis.** *Comput Methods Programs Biomed* 2011, **102:**156-171.
30. Langouche L, Vander Perre S, Wouters PJ, D'hoore A, Hansen TK, Van den Berghe G: **Effect of intensive insulin therapy on insulin sensitivity in the critically ill.** *The Journal of Clinical Endocrinology & Metabolism* 2007, **92:**3890-3897.
31. Thomas F, Pretty CG, Fisk L, Shaw GM, Chase JG, Desaive T: **Reducing the impact of insulin sensitivity variability on glycaemic outcomes using separate stochastic models within the STAR glycaemic protocol.** *Biomedical engineering online* 2014, **13:**43.
32. Lin J, Razak NN, Pretty CG, Le Compte A, Docherty P, Parente JD, Shaw GM, Hann CE, Chase JG: **A physiological Intensive Control Insulin-Nutrition-Glucose**

(ICING) model validated in critically ill patients. *Comput Methods Programs Biomed* 2011, **102:**192-205.
33. Chase JG, Suhaimi F, Penning S, Preiser J-C, Le Compte AJ, Lin J, Pretty CG, Shaw GM, Moorhead KT, Desaive T: **Validation of a model-based virtual trials method for tight glycemic control in intensive care.** *Biomedical engineering online* 2010, **9:**84.
34. Lin J, Lee D, Chase JG, Shaw GM, Hann CE, Lotz T, Wong J: **Stochastic modelling of insulin sensitivity variability in critical care.** *Biomedical Signal Processing and Control* 2006, **1:**229-242.
35. Lin J, Lee D, Chase JG, Shaw GM, Le Compte A, Lotz T, Wong J, Lonergan T, Hann CE: **Stochastic modelling of insulin sensitivity and adaptive glycemic control for critical care.** *Comput Methods Programs Biomed* 2008, **89:**141-152.
36. Stewart KW, Pretty CG, Tomlinson H, Thomas FL, Homlok J, Noémi SN, Illyés A, Shaw GM, Benyó B, Chase JG: **Safety, efficacy and clinical generalization of the STAR protocol: a retrospective analysis.** *Annals of intensive care* 2016, **6:**24.
37. Uyttendaele V, Dickson J, Shaw G, Desaive T, Chase J: **Improved Blood Glucose Forecasting Models using Changes in Insulin Sensitivity in Intensive Care Patients.** 2017.
38. Uyttendaele V, Knopp JL, Stewart KW, Desaive T, Benyó B, Szabó-Némedi N, Illyés A, Shaw GM, Chase JG: **A 3D insulin sensitivity prediction model enables more patient-specific prediction and model-based glycaemic control.** *Biomedical Signal Processing and Control* 2018, **46:**192-200.
39. Dickson JL, Stewart KW, Pretty CG, Flechet M, Desaive T, Penning S, Lambermont BC, Benyó B, Shaw GM, Chase JG: **Generalisability of a virtual trials method for glycaemic control in intensive care.** *IEEE Trans Biomed Eng* 2018, **65:**1543-1553.
40. Chase JG, Preiser J-C, Dickson JL, Pironet A, Chiew YS, Pretty CG, Shaw GM, Benyo B, Moeller K, Safaei S: **Next-generation, personalised, model-based critical care medicine: a state-of-the art review of in silico virtual patient models, methods, and cohorts, and how to validation them.** *Biomedical engineering online* 2018, **17:**24.
41. Stewart KW, Pretty CG, Tomlinson H, Fisk L, Shaw GM, Chase JG: **Stochastic model predictive (stomp) glycaemic control for the intensive care unit: development and virtual trial validation.** *Biomedical Signal Processing and Control* 2015, **16:**61-67.
42. Docherty P, Chase J, David T: **Characterisation of the iterative integral parameter identification method.** *Med Biol Eng Comput* 2012, **50:**127-134.
43. Hann CE, Chase JG, Shaw GM: **Integral-based identification of patient specific parameters for a minimal cardiac model.** *Comput Methods Programs Biomed* 2006, **81:**181-192.
44. Wong X, Singh-Levett I, Hollingsworth L, Shaw G, Hann C, Lotz T, Lin J, Wong O, Chase J: **A novel, model-based insulin and nutrition delivery controller for glycemic regulation in critically ill patients.** *Diabetes Technol Ther* 2006, **8:**174-190.
45. Docherty P, Chase JG, Lotz T, Desaive T: **A graphical method for practical and informative identifiability analyses of physiological models: A case study of insulin kinetics and sensitivity.** *BioMedical Engineering OnLine* 2011, **10:**39.
46. Sheather SJ: **Density estimation.** *Statistical science* 2004**:**588-597.
47. Ghasemi A, Zahediasl S: **Normality tests for statistical analysis: a guide for non-statisticians.** *International journal of endocrinology and metabolism* 2012, **10:**486.
48. Kohavi R: **A study of cross-validation and bootstrap for accuracy estimation and model selection.** In *Ijcai*. Montreal, Canada; 1995: 1137-1145.

Appendix A

Silverman's rule of thumb is used to calculate a local kernel standard deviation, σ_{loc}, from a global standard deviation, σ_{gbl}. For k data point and n dimensions, Silverman's rule of thumb can be written [46]:

$$\sigma_{\text{loc}} = \left(\frac{4}{n+2}\right)^{\frac{1}{n+4}} \times k^{-\frac{1}{n+4}} \times \sigma_{\text{gbl}} \tag{A1}$$

Eq. A1 can be modified to give a kernel specific local standard deviation using a relative local density factor as in [34]:

$$\sigma_{\text{loc}} = \left(\frac{4}{n+2}\right)^{\frac{1}{n+4}} \times \left(\frac{\rho_{\text{loc}}}{\rho_{\text{gbl}}} k\right)^{-\frac{1}{n+4}} \times \sigma_{\text{gbl}} \tag{A2}$$

where ρ_{loc} is the kernel data point density, in ortho-normalised space, within a radius $k^{-1/7}$ (for the 3D case) from a given kernel data point, and ρ_{gbl} is the global kernel data point density, again in ortho-normalised space. For the 3D case:

$$\rho_{\text{loc}} = \frac{m}{\frac{4}{3}\pi\left(k^{-\frac{1}{7}}\right)^3} \tag{A3a}$$

$$\rho_{\text{gbl}} = \frac{Zk}{\frac{4}{3}\pi {R_Z}^3} \tag{A3b}$$

$$\frac{\rho_{\text{loc}}}{\rho_{\text{gbl}}} = \frac{m {R_Z}^3 k^{\frac{3}{7}}}{Zk} \tag{A4}$$

where m is the number of ortho-normalised kernel data points that lie within a radius $k^{-1/7}$ from a given kernel data point, and R_Z is defined as the radius from the global ortho-normalised centre point of 0 that contains Zk kernel data points for some $Z \leq 1$.

Substituting Eq. A4 into Eq. A2 with $n = 3$ yields:

$$\sigma_{\mathrm{loc}} = 0.969 \times \left(\frac{m{R_Z}^3 k^{\frac{3}{7}}}{Z} \right)^{-\frac{1}{7}} \times \sigma_{\mathrm{gbl}} \tag{A5}$$

And as, in ortho-normalised space, $\tilde{\sigma}_{\mathrm{gbl}}$ is 1 in every direction, Eq. A5 can be further simplified:

$$\tilde{\sigma}_{\mathrm{loc}} = 0.969 \times \left(\frac{m{R_Z}^3 k^{\frac{3}{7}}}{Z} \right)^{-\frac{1}{7}} \tag{A6}$$

As in Eq. 13.

Appendix B

Working in ortho-normalised space, for a given transformed 2D sample point $\tilde{\mathbf{y}}_{1:2}$, the conditional 1D PDF is exists along a line, $\tilde{\mathbf{Y}}$, defined by varying $\tilde{y}_3$. This behaviour is consistent with the non-transformed case where $\mathbf{Y}$ is defined by varying y_3 due to the transformation matrix $\mathbf{A}$ being upper triangular i.e. only $\tilde{y}_3$ varies as y_3 is varied. The goal is to define the region on the line $\tilde{\mathbf{Y}}$ where the distance from a given kernel centre point, $\tilde{\mathbf{x}}_j$, to the sample point, $\tilde{\mathbf{y}}$, is less than 5 kernel standard deviations, $\tilde{\sigma}_j$:

$$\left\|\tilde{\mathbf{y}} - \tilde{\mathbf{x}}_j\right\|_2 < 5\tilde{\sigma}_j \tag{B1}$$

The minimum possible distance between the line $\tilde{\mathbf{Y}}$ and the kernel centre point $\tilde{\mathbf{x}}_j$ occurs when $\tilde{y}_3 = x_{j,3}$, and the norm reduces to the norm of the first 2, fixed components of each:

$$\left\|\tilde{\mathbf{y}}_{1:2} - \tilde{\mathbf{x}}_{j,1:2}\right\|_2 < 5\tilde{\sigma}_j \tag{B2}$$

If the condition in B2 is not satisfied, the current kernel centre point has almost no effect on the PDF along the line $\tilde{\mathbf{Y}}$, and the method iterates immediately to the next kernel centre point. If the condition in B2 is satisfied, iteration now occurs across the set of points on $\tilde{\mathbf{Y}}$ within a radius $5\tilde{\sigma}_j$ from $\tilde{\mathbf{x}}_j$, defined by restricting the range of $\tilde{y}_3$:

$$\tilde{x}_{j,3} - \sqrt{\left(5\tilde{\sigma}_j\right)^2 - \left(\tilde{\mathbf{y}}_{1:2} - \tilde{\mathbf{x}}_{j,1:2}\right)^2} \geq \tilde{y}_3 \geq \tilde{x}_{j,3} + \sqrt{\left(5\tilde{\sigma}_j\right)^2 - \left(\tilde{\mathbf{y}}_{1:2} - \tilde{\mathbf{x}}_{j,1:2}\right)^2} \tag{B3}$$

and iterating between these minimum and maximum values for $\tilde{y}_3$, rather than across the full possible range of values, unless the existing minimum or maximum values are already more restrictive. Note that these bounds are specific to both the sample line, $\tilde{\mathbf{Y}}$, and kernel point, $\tilde{\mathbf{x}}_j$.